

# A Novel Development of Acoustic SLAM

Joseph O'Reilly

*Science & Engineering. Anglia Ruskin  
University*  
Cambridge, Cambridgeshire  
joseph.oreilly@pgr.anglia.ac.uk

Silvia Cirstea

*Science & Engineering. Anglia Ruskin  
University*  
Member, IEEE  
Cambridge, Cambridgeshire  
silvia.cirstea@anglia.ac.uk

Marcian Cirstea

*Science & Engineering. Anglia Ruskin  
University*  
Senior Member, IEEE  
Cambridge, Cambridgeshire  
marcian.cirstea@anglia.ac.uk

Jin Zhang

*Science & Engineering. Anglia Ruskin  
University*  
Member, IET  
Cambridge, Cambridgeshire  
jin.zhang@anglia.ac.uk

**Abstract**— This paper will explore and develop on the novel idea of using acoustics to map and navigate indoor environments. The system requirements, modelling and evaluation are addressed, alongside the design and development process, testing methods, desired outcomes and practical applications. Previous work carried out in this field demonstrates that it is possible to use first order echoes to map a room. The current paper is reporting on initial research to further develop such algorithms into a simultaneous localization and mapping algorithm, having the capability to not only map rooms with sound but to also navigate rooms as well. Such novel system is intended to help visually impaired people to navigate rooms by making use of sounds and their echoes, thus 'listening' their way into navigating through a room. The paper overviews the approach taken towards developing a navigation algorithm using sound, as well as the associated modelling, simulation and testing strategies enabling the desired outcomes of this type of system.

**Index Terms**— Kalman filters, Autonomous systems, Acoustical engineering, Navigation, Simultaneous localization and mapping, Acoustic signal processing.

## I. INTRODUCTION

THIS paper covers aspects related to the development process, testing methods, desired outcomes and final uses for the novel idea of using acoustics to map and navigate indoor environments. Some interesting questions to be addressed relate to the ability of autonomous navigation and mapping to be achieved using only audible sound. If so, another question to be answered is if this can be done by only using environmental noise.

It is known that some visually impaired people are capable of navigating rooms by listening to the echoes created by clicking their fingers, so, naturally one wonders if an intelligent system dedicated for such purpose (a robot) could do the same, thus helping the visually impaired to navigate.

A survey of previous work done shows that an interesting question was posed: "Can one hear the shape of a drum?" [1]; the research determined that drums of different shape don't necessarily resonate at different frequencies. Other works carried out [2] developed this theory and applied it to an indoor room, demonstrating that it is possible to use first order echoes to map a room. The present work follows up in further developing this algorithm into a simultaneous localization and mapping algorithms enabling the capability to not only map rooms with sound but also to navigate rooms as well.

The interaction between ambient sound and objects in the environment, manifested through reverberation and echoes, is an important source of information about the composition of the scene, insufficiently exploited in autonomous navigation systems, which can solve the problem of navigation in low light conditions. Show the application of acoustic imaging for mapping a stationary room using a distributed array of five microphones which capture the room impulse response (RIR). This contains the times of arrival of the reverberations from walls and other features using an artificial intelligence (AI) techniques called multidimensional scaling (MDS). This demonstrates that it has been possible to estimate the most likely association between a given reverberation and a particular wall and thus reconstruct the map of the room.

Research in the area of autonomous navigation has developed complex AI systems that combine sensors like cameras, Light Imaging Detection And Ranging (LIDAR) and Inertial Measuring Units (IMUs) with Simultaneous Localization and Mapping (SLAM) algorithms [3] SLAM was developed to guide robots [4] by estimating the map of an unknown environment and, at the same time, updating the position of the robot in relation to this environment, based on probabilistic estimation of feature points (e.g. walls, obstacles) from time-varying inputs of exteroceptive sensors attached to the robot [5]. Sensors like LIDAR and cameras have the advantage of being accurate and of high resolution, however, they come with their downsides: LIDAR is a very expensive sensor and poses health and safety problems in operation; cameras, while becoming less costly, require high processing power as well as having low signal to noise ratios in low light environments.

By using acoustic imaging techniques and a SLAM algorithm adapted for feature detection from sound for navigation in a low light environment, a novel navigation system can be developed. This would have the potential to be used as an assisted living aid for the visually impaired or in robotic navigation for fire-fighting and exploration of dark spaces.

In this paper Section II and III focus on the two major research areas of this novel development. Section II covers the methods of gathering information from the environment with sound and how this information is used to map the environment. Section III overviews the different methodologies and algorithms used in SLAM systems with a focus on indoor environments using microphones as the primary sensor. The development process, testing breakdown and more complex developments is shown in section IV, with the preliminary testing results of RIR input methods is shown in section V.

## II. ACOUSTIC MAPPING

Sound holds a lot of information of its environment, enough information to accurately estimate the shape of a room. Dokmanic et al.[2] has proved this by developing an algorithm that uses the information from a RIR recorded by an array of microphones to locate the walls in which the echoes from the RIR reflected off of, thus mapping the room. This section will highlight the key elements of RIR and Dokmanic's algorithm.

### A. Room impulse response

An impulse response is one of the most important tests one could use to learn the characteristics of a linear system. For this purpose, the room is a linear system. This is where the output of a system is recorded when a delta function is applied at the input. The impulse response of a room is commonly used in acoustic analysis to find acoustic "dead spots", early to late ratios and reverberation time. The reverberation characteristic of a room is commonly used to computationally recreate reverberation for musical application but for acoustic mapping this is the main characteristic which enables the location of the walls of the environment.

There are multiple methods for gathering a RIR: Impulse, Chirp, Linear/Logarithmic sweep, Maximum-Length Signal (MLS) and Inverse Repeated Signal (IRS). These have been compared and tested in many papers [6], [7] and [8]. However, the testing parameters from these works have never explored the situation where the receiver, source or environment is moving. This is due to it not being necessary for a typical RIR use case but for acoustic SLAM, knowing how these methods are affected by a dynamic situation is of high importance.

Impulse is the standard linear system input for an impulse response. The problem with using this method for RIR is that it generates very little energy, which is necessary to acquire clear reverberation in the output. To combat this, it is common to use an impulse train, with multiple impulses, and then calculate the average. Due to these issues, impulse input is not generally used for RIR.

Chirp generates a linear frequency sweep over a short period of time (10-100ms) which is then repeated every 2-3 seconds. By using a range of frequencies it reduces the possibility of results being effected by the frequency responses of different rooms. By fast Fourier transform (FFT) deconvolution of the input from the output the RIR can be calculated. This is done by taking the FFT of the input divide it by the FFT of the output and then taking the inverse fast Fourier transform (IFFT) of the result.

Linear/Logarithmic sweep is like a chirp but instead of an impulse train, a longer continuous signal is used. The main advantage of this method over a chirp is that more power is being produced, which in turn results in a better signal to noise ratio.

MLS is close to a white noise signal that has a constant magnitude and a uniform random phase spectra. By using a MLS, the RIR can be found with circular cross-correlation between the output and MLS input, as the generated signal will be unique in its environment. MLS has a long history as a primary input method for RIR as it has many advantages,

which are described in [7]. One issue of MLS is that it can produce phantom echoes when faced with non-linear distortion. This issue could be catastrophic, as the primary purpose of the RIR in acoustic mapping is to find the echoes. IRS uses two MLS sequences, one positive and one negative, and alternates between the two. In doing so it doubles the measurement time but increases the immunity to distortion.

### B. Mapping a room with sound

Dokmanic et al's.[2] method for mapping the walls of a room is to use an array of 4 or more microphones to record the RIR from a sweeping sine wave generated by a source with an unknown location. The initial audio impulse and all echoes from the RIR are labeled with their time of arrival and magnitude for each individual microphone in the array. The echo information is compared between each microphone to find the direction each echo came from. The labelling process is carried out by taking advantage of Euclidean Distance Matrices (EDM) arithmetic's: each microphone's distance of placement apart from each other is placed in an EDM. The pairing impulses can then be found by entering the distance between the microphone and each recorded impulse virtual source into a new row of the EDM and then by using EDM arithmetic's, false combinations can be found.

Once the directions of all the echoes are known, the room's shape can be constructed. When noise is factored in, there can be a situation where no combinations of the echoes source creates an EDM. In this situation, the probability of how close each arrangement is to a true EDM value is calculated. For a single room measurement, the one which is closest to a true value is used, but in regards to using this method with SLAM it will be beneficial to use a statistical filter for the probability of each EDM arrangement which updates for each sample. Statistical filters are commonly used in SLAM for each sample of sensor data but in this unique situation the sensor information could have multiple outcomes per sample; this will be discussed further in Section III and IV.

It has been shown in Dokmanic's work that this is a reliable method for mapping the walls in a room due to its high precision without even needing to know the location of the source sound or the receiver locations. This means that in a practical environment where the system is constantly moving, recalibration of the receivers can be done purely in the software. This also leads onto another possible improvement where multiple sources from different locations in the environment could be layered on top of each other, allowing a system to not only need to generate a sound to produce a map but also compare multiple generated maps to find the highest probability of creating an accurate map.

## III. SIMULTANEOUS LOCALIZATION AND MAPPING

For truly autonomous robotic navigation to be achieved, when an unknown environment is presented, with an unknown position in this environment, a robot must be capable of mapping its soundings with its onboard sensors and accurately estimating its location on this map. This is known as the SLAM problem statement. This allows a robot to be deployed in situations where the area is unknown and the user cannot see the robot, such as fire rescue in high smoke density buildings, small tunneled cave systems and assistance for the visually impaired.

Due to the nature of robotics, different situations require different solutions for navigation: the environment, (indoors, outdoors, underwater, airborne) the sensors used (laser, camera, IMU, audio) and the moving system (wheeled, flying, walking, etc.) all affect the SLAM algorithm.

The scope of this paper is focused on the mapping of the environment before it is used with SLAM. So, this paper gives an overview of SLAM methods which are relevant to an indoor environment using microphones as the primary<sup>1)</sup> input sensor for gathering information of the environment.

#### A. General Methodology

The main three elements of SLAM are: the map representation, the data processing and the sensors used. Each of these building blocks are essential for the development of a SLAM algorithm. For this system, the primary sensor will be an array of microphones that gathers the location of a room's walls through sound. The main development process for acoustic SLAM will be finding the best fit of map representation and data processing.

#### B. Mapping Methods

The mapping method is how the SLAM algorithm information is visually processed. There are many methods to do this: Topological, Semantic, Appearance and Hybrid which have been compared and expanded upon in various works [9], [10]. Due to the vast information of these methods this paper covers the main two methods: 1) feature mapping and 2) grid mapping.

##### 1) Feature mapping

The standard model observation method used in SLAM is landmark/feature detection. This involves labeling landmarks in the environment and using them as reference points for calculating the position of the system. Normally, this method is used with camera based systems, where landmarks in the environment are easily determined with shape, color and depth.

The more unique features that can be recorded, the higher the accuracy of the map. For range-based sensors (laser, ultrasonic, echo), only edge detection and free standing objects can be used as landmarks, which results in clear landmarks but usually less of them compared to cameras.

##### 2) Grid mapping

Rather than mapping features and landmark locations grid based SLAM uses grid maps. This approach splits the environment into ridged independent cells which are recorded as either occupied space or free space. This generates a floor plan like map, where occupied cells are black pixels and free space is white. The main disadvantage of grid based SLAM is that computationally it takes up more memory as the entire map needs to be recorded, compared to feature based SLAM where only the feature locations needs to be recorded. The main advantage of this method though is that feature detection is not required, meaning the system can work directly with the raw sensor data. This is preferable for niche sensors methods where the amount of features are limited.

#### C. Data Processing methods

The data processing methods refer to how the system calculates its most probable location in the map. The two most common methods are Bayesian based filters and partial based filters, both methods being quite old. There are also some interesting new AI based method which aim to overcome some of the major issues the older methods have.

##### Bayesian filter framework

For feature mapping techniques, it is common to use a Bayesian filter framework for state estimation. Bayes' theorem describes the probability of an event, based on prior knowledge of conditions; this creates the framework for state estimation.

This breaks down the SLAM problem in to its probabilistic form, by calculating the probability of the system's current location (defined as ' $x_k$ ' where  $x$  is the locations vector and  $k$  is the time period) when all observed landmarks (defined as ' $z_{1:k}$ ', where ' $z$ ' is an array of landmark location vectors) and movements of the system (defined as ' $u_{1:k}$ ', where ' $u_k$ ' is the system's current movement vector) are known.

$$bel(x_k) = P(x_k | z_{1:k}, u_{1:k}) \quad (1)$$

This can be defined by a simple probability equation, where the current location is dependent on all past and current observed landmarks and all past and current vector movements of the system. Using statistical methods, shown in [11], the generic SLAM algorithm can be represented through two equations: the prediction step and the correction step.

The prediction step (2) calculates the probability of the systems next possession based on the previous position of the system. This requires the motion model, (3) which models the uncertainty in the optometry data by finding the probability of the current position of the system in relation to the recorded distance traveled and the previous location of the system.

$$\overline{bel}(x_k) = \int P(x_k | x_{k-1}, u_k) bel(x_{k-1}) dx_{k-1} \quad (2)$$

$$P(x_k | x_{k-1}, u_k) \quad (3)$$

The correction step (4) calculates the probability of the systems current possession based on the sensor data of the environment. This requires the observation model (5), which calculates the probability of a landmark being observed in relation to where the system currently believes it is based on the motion model.

$$bel(x_k) = \eta P(z_k | x_k) \overline{bel}(x_k) \quad (4)$$

$$P(z_k | x_k) \quad (5)$$

The most common filters this method is based on are the Kalman filter (KF) [11] and the less computationally intensive Extended Kalman filter (EKF) [12]. KF-SLAM is the optimal estimator when two assumptions are made: all models are linear and all distributions are Gaussian. EKF-SLAM is highly influenced by incorrect observations of landmarks. This is often a problem when the vehicle returns

to a previously visited location after traveling a large distance, this is known as the “loop-closure” problem.

## 2) *Particle Filter Framework*

Particle filters represent a probability distribution as a line of particles (samples) where the density of the particles represents the probability. When there is a new observation, the points are weighted by the observation probability, then the distribution is resampled and a new distribution is made. Fast-SLAM [13] is currently one of the most used SLAM algorithms, which is favored over the EKF-SLAM one due to the fact that the filter allows non-linear distributions, making it less restrictive when used in a realistic environment.

## 3) *Other Filter Frameworks*

The main issue with KF-SLAM and Fast-SLAM is that they assume the environment is static, which is not realistic. Novel approaches to overcome this issue involve using neural networks and AI, one of them being generalized motion (GEM)-SLAM [14]. GEM-SLAM shows a large improvement in accuracy when there are false readings and uncertainty caused by dynamic objects in the environment.

# IV. DEVELOPMENT OF A NOVEL ACOUSTIC SLAM METHOD

This paper presents aspects of the development for an autonomous echo-location system based on acoustic imaging and acoustic SLAM. Such systems can be used for guided navigation in dynamic scenes affected by environmental noise.

## A. *Stage 1: Initial developments*

There are two interconnecting challenges in the initial development of this novel mapping system: I) the acoustic mapping technique has to produce the highest quality environment information for the SLAM algorithm, and II) the SLAM algorithm being used must take as much advantage as possible of the provided information set. Even though these aspects are dependent on each other, they can be addressed independently due to how they are interconnected. As the initial development will be using Dokmanic’s algorithm [2], the environment data being handed over to the SLAM algorithm will be in a constant format even if the acoustic recording methods change.

As shown in section II, there are many different input methods for a RIR and for a robust system the one which gives the clearest reverberation whilst having an acceptable signal to noise ratio (SNR) in a realistic environment needs to be found. Using an array of microphones, a range of RIR input methods will be tested. Room shape, room size, input volume, environmental noise and level of clutter are the main elements which need to be tested and how well each method does will be based on the SNR of the output and the accuracy of mapping. When the optimal RIR method is found/developed, there are multiple situations that need to be tested which go beyond the scope of Dokmanic’s work. Useful information like environment material (wood, glass, carpet), open and closed doors or specific object detection could be found inside the RIR which would be useful to take advantage of if the computational cost is not too high.

Dokmanic’s algorithm creates a 3D graph with the walls, source and receivers plotted as vectors. It would then seem

almost natural to use a graph or grid based method but then, as a 3D map will be developed, the computational costs of these methods greatly increase. Feature based SLAM is less costly but the lack of information may result in unreliable mapping. The SLAM algorithm for the initial testing will need to accurately map with the information from the acoustic mapping algorithm and also be low powered enough to be implemented onto a portable microcontroller. The filter and mapping technique both need to be explored and tested to see which combination produces the most accurate map whilst using the least power.

The acoustic method and SLAM algorithm from these tests will develop the initial system where they will be implemented into a standalone hardware board which will control a wheeled robot. At this point the system will be capable enough to develop a map of its environment with echolocation and when the system is moved it will be able to locate itself. At this point, the system was only been tested by gathering information stationary but in a realistic setting it is assumed that a navigation robot is capable of SLAM whilst moving. This is the next step of development.

## B. *Stage 2: System developed for realistic environments*

The previously mentioned acoustic mapping methods have only been tested using a stationary environment. Sound acts differently when its source is moving and when it interacts with moving objects, this is known as the Doppler Effect. This means that the previously tested RIR input methods, which gave an acceptable stationary result, need to be tested whilst the source and receivers are moving at different speeds, directions and rotations. In turn, this means that an IMU needs to be implemented onto the standalone hardware board to accurately measure the movement of the system. The change in accuracy for these movement elements will be used to develop a margin of error that the SLAM algorithm can use for the probability filter.

Once the system is at a point where it can accurately map whilst it is moving, it will then need to be able to navigate a realistic environment where there are moving objects within it. Firstly, this stage is dependent on how the current system is affected by moving objects, as unlike other sensors the echo location algorithm only maps walls. The next investigation stage would be to test how the system is affected by different size and speeds of moving objects whilst the system is navigating.

If the system is noticeably affected, then there are two development approaches that can be made to solve this problem: develop the acoustic mapping to detect and filter moving objects, or use a more complex SLAM algorithm that can support a dynamic environment. As stated in section III, the majority of SLAM algorithms assume the environment is stationary; this is not realistic. There are SLAM algorithms that tackle this issue [16], though more complex algorithms require higher computational power. The testing process will be repeated from stage 1 for these more complex SLAM algorithms compared with the already functioning algorithm and the further developed algorithm which filters moving objects. This test compares the accuracy and power consumption in a dynamic environment.

### C. Stage 3: Environmental sound mapping

An interesting factor of previous works [2] is that the sound source location is not need to locate the walls. A sound created in the environment naturally could be used to map the room if it can be successfully isolated. This would mean that the system could navigate purely based on environmental noises being used as inputs for RIR, applying no additional impact into its environment. If a single environmental sound can successfully map a room, it could also be possible to use multiple RIR each using different sounds being computed in parallel, which in theory would increase the accuracy of mapping.

The major problem with this theory is that all methods of RIR use deconvolution, which requires the input signal to be known. This means a new RIR method will need to be developed, which does not have any prior knowledge of the impulse or can estimate what the input signal is before deconvolution.

## V. INITIAL EVALUATION AND TESTING OF THE METHOD

### A. Evaluation principles

The first stage of testing compares commonly used RIR input methods (impulse, chirp, logarithmic sweep, MLS and popping a balloon) in regards to their SNR in different indoor room environments (Fig 2.) and at different input volumes. Each of the RIR, except popping of a balloon, was generated in Matlab and transmitted by using a Digital Audio Workstation (DAW). The DAW also recorded the four microphones in sync and exported the signal information to be manipulated in Matlab.

For clear results, all initial testing was done with high quality hardware, before the system is developed into a lower power

system. An array of 4 microphones was used as audio sensor to record the RIR. It is important that the microphones used are capable of omnidirectional recording to allow no bias to any direction of sound. High-quality microphones require a high amount of power, referred to as phantom power, which is usually provided by the audio interface. The microphones used in these tests were four AKG C414-XLS condenser microphones, used in omnidirectional mode. When using high quality microphones, they often require in line power, amplification and to be converted from analogue to digital to be processed by a computer, which the standard input/output (I/O) ports of a computer cannot provide. Audio interfaces power the microphones and convert their analogue signal into digital for a computer to take advantage of it. The individual important factors of the audio interface are the amount of microphone inputs and the comparability and quality of the audio drivers it can use. The Audio interface used was the ALESIS MULTIMIX 8 USB FX, which has four microphone inputs and uses a common high quality audio driver audio stream input/output (ASIO). A directional Yamaha loud speaker produces the RIR input signal. Both directional speakers and omnidirectional speakers can be used but omnidirectional speakers produce a more realistic spreading sound due to sending sound in all directions. Although these types of speakers are expensive and not necessary for the preliminary testing, both the microphones, input of the audio interface and output of the audio interface all have amplifier settings which are all kept at 0dB.

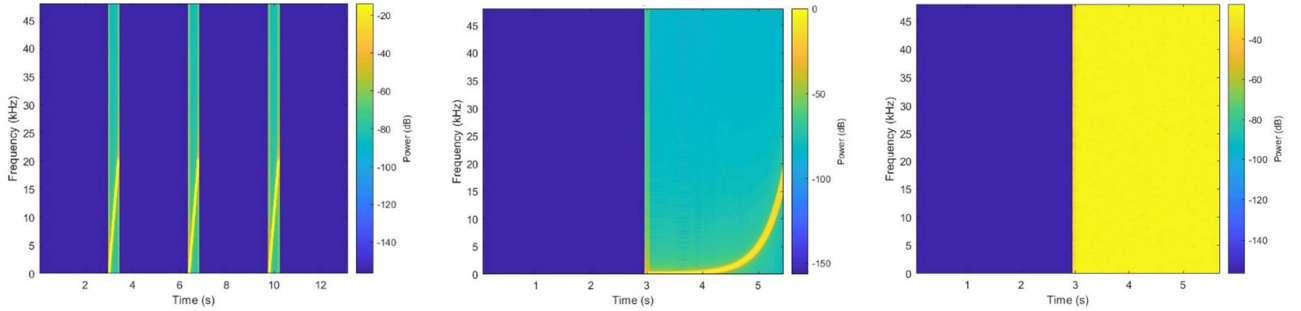


Fig 1. 20Hz to 20kHz linear chirp train, 20Hz to 20kHz logarithmic sweep and MLS RIR input signals used for testing.



Fig 2. Three of the rooms used for testing, each with different shapes and volumes of clutter.





## B. Test results

The RIR signals generated in Matlab (Fig.1) were tested and recorded in three separate rooms (Fig.2). The RIR for each room is shown in Fig.3. From these results there are some clear observations that can be made. The low input power limitation of the impulse signal is clearly visible. The signal output is lower by a magnitude of 10 meaning the ambient noise of the system is noticeable compared to the other methods.

Fig 3 recordings is when there is no reduction in the power of the input signal. These methods were also tested at -12dB and -30dB, which was applied at the DAW output. As the signal power is reduced the SNR of each RIR decreases as expected, at -30dB the impulse RIR error rate is too high to obtain any useful information.

In respect for the system to function without knowledge of the input signal the MLS method seems unlikely to be useful to develop upon. This is due to the deconvolution being impossible without previous knowledge of the unique signal.

## VI. CONCLUSION

Autonomous robotic navigation is an ever developing field, constantly evolving and improving on itself. There is a vast difference in environments, where robots are required to navigate, certain environments requiring specific solutions. Acoustic SLAM takes advantage of the amount of information the sound holds about its environment which is an important tool when designing navigation systems.

The approach taken towards developing a novel navigation algorithm using sound has been covered in this paper, with an overview of the two main elements required to develop it, acoustic mapping and SLAM, and the testing methodology that is being undergone to develop it. With the inclusion of the preliminary testing results of the RIR input in different environments.

When finalized, the completed system would be capable of navigating without needing to generate a sound, creating no impact in the environment. This means that this system will be able to function alongside people without any risk. Not only would this system be safe but it would also be capable to map a 360° three-dimensional room with a single microphone array, which no other sensor that is currently being used with SLAM can do.

In the longer term, and beyond the scope of this project, the concept should lead to a practical solution that would be easy to use and require little or no external infrastructure, to make it ubiquitous available. In addition, such a solution should have minimal impact on the surroundings, be environmentally friendly, robust and cost-effective (both to acquire and use). Other further developments would include integrating acoustic SLAM with other commonly used sensors to increase accuracy for generalized robotic navigation. Sound holds a lot of information about the environment which can be used to support LIDAR and camera based systems, providing more information at a low system cost.

## REFERENCES

- [1] K. Mac, "Can't one really hear the shape of a drum?," vol. 57, no. 4, p. 465, Jul. 1966.
- [2] I. Dokmanić, R. Parhizkar, A. Walther, Y. M. Lu, and M. Vetterli, "Acoustic echoes reveal room shape," vol. 110, no. 30, pp. 12186–12191, Jul. 2013.
- [3] Evers C, Naylor PA. "Acoustic SLAM." *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 26(9):1484–98, Sep. 2018;
- [4] Dissanayake MG, Newman P, Clark S, Durrant-Whyte HF, Csorba M. "A solution to the simultaneous localization and map building (SLAM) problem." *IEEE Transactions on robotics and automation*. 17(3):229–41 Jun. 2001.
- [5] A. J. Davison, "Mobile Robot Navigation Using Active Vision," Ph.D. dissertation, Dept. Science Eng., Oxford Univ., Oxford, UK, 1998.
- [6] A. Carini, S. Cecchi, L. Romoli, A. Carini, S. Cecchi, and L. Romoli, "Robust Room Impulse Response Measurement Using Perfect Sequences for Legendre Nonlinear Filters," vol. 24, no. 11, pp. 1969–1982, Nov. 2016.
- [7] Stan GB, Embrechts JJ, Archambeau D. "Comparison of different impulse response measurement techniques." *Journal of the Audio Engineering Society*. 50(4):249–62. 2002;
- [8] Farina A. "Simultaneous measurement of impulse response and distortion with a swept-sine technique." *In Audio Engineering Society Convention*, 108, Feb. 2000.
- [9] Canclini A, Antonacci F, Sarti A, Tubaro S. "Acoustic source localization with distributed asynchronous microphone networks." *IEEE Transactions on Audio, Speech, and Language Processing*. 21(2):439–43. Feb 2013.
- [10] Norcross S, Bradley JS. "Comparison of room impulse response measurement methods." *Canadian Acoustics*. Sep. 1994.
- [11] Welch G, Bishop G. "An introduction to the Kalman filter." 1995.
- [12] Ljung L. "Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems." *IEEE Transactions on Automatic Control*. Feb. 1979.
- [13] Montemerlo M, Thrun S, Koller D, Wegbreit B. "FastSLAM: A factored solution to the simultaneous localization and mapping problem." Jul. 2002.
- [14] C. Evers and P. Naylor, "Optimized Self-Localization for SLAM in Dynamic Scenes Using Probability Hypothesis Density Filters," vol. 26, no. 9, pp. 1484–1498, Sep. 2018.